

Site-specific temporal and spatial validation of a generic plant pest forecast system with observations of *Bactrocera dorsalis* (oriental fruit fly)

Seung Cheon Hong¹, Roger D. Magarey², Daniel M. Borchert³,
Roger I. Vargas², Steven K. Souder⁴

1 Department of Plant Pathology, North Carolina State University, Raleigh, North Carolina, USA **2** North Carolina State University and Cooperator with USDA-APHIS-PPQ-CPHST-PEARL, 1730 Varsity Drive, Suite 300, Raleigh, North Carolina, USA **3** Animal and Plant Health Inspection Service-Plant Protection and Quarantine-Center for Plant Health Science and Technology, Plant Epidemiology and Risk Analysis Laboratory, Raleigh, North Carolina, USA **4** United States Department of Agriculture, Agricultural Research Services, U.S. Pacific Basin Agricultural Research Center, Hilo, Hawaii, USA

Corresponding author: Roger D. Magarey (rdmagare@ncsu.edu)

Academic editor: Richard Shaw | Received 23 April 2015 | Accepted 13 August 2015 | Published 15 September 2015

Citation: Hong SC, Magarey RD, Borchert DM, Vargas RI, Souder SK (2015) Site-specific temporal and spatial validation of a generic plant pest forecast system with observations of *Bactrocera dorsalis* (oriental fruit fly). NeoBiota 27: 37–67. doi: 10.3897/neobiota.27.5177

Abstract

This study introduces a simple generic model, the Generic Pest Forecast System (GPFS), for simulating the relative populations of non-indigenous arthropod pests in space and time. The model was designed to calculate the population index or relative population using hourly weather data as influenced by developmental rate, high and low temperature mortalities and wet soil moisture mortality. Each module contains biological parameters derived from controlled experiments. The hourly weather data used for the model inputs were obtained from the National Center of Environmental Prediction Climate Forecast System Reanalysis (NCEP-CFSR) at a 38 km spatial resolution. A combination of spatial and site-specific temporal data was used to validate the GPFS models. The oriental fruit fly, *Bactrocera dorsalis* (Hendel), was selected as a case study for this research because it is climatically driven and a major pest of fruit production. Results from the GPFS model were compared with field *B. dorsalis* survey data in three locations: 1) Bangalore, India; 2) Hawaii, USA; and 3) Wuhan, China. The GPFS captured the initial outbreaks and major population peaks of *B. dorsalis* reasonably well, although agreement varied between sites. An index of agreement test indicated that GPFS model simulations matched with field *B. dorsalis* observation data with a range between 0.50 and 0.94 (1.0 as a perfect match). Of the three locations, Wuhan showed

the highest match between the observed and simulated *B. dorsalis* populations, with indices of agreement of 0.85. The site-specific temporal comparisons implied that the GPFS model is informative for prediction of relative abundance. Spatial results from the GPFS model were also compared with 161 published observations of *B. dorsalis* distribution, mostly from East Asia. Since parameters for pupal overwintering and survival were unknown from the literature, these were inferred from the distribution data. The study showed that GPFS has promise for estimating suitable areas for *B. dorsalis* establishment and potentially other non-indigenous pests. It is concluded that calibrating prediction models with both spatial and site-specific temporal data may provide more robust and reliable results than validations with either data set alone.

Keywords

Risk analysis, invasive species, modeling, climate

Introduction

The increase in international trade has exacerbated the problem of non-indigenous species moving between continents (Lenzen et al. 2012) and causing economic damage. Phytosanitary regulatory agencies aim to prevent non-indigenous pest entry and establishment and attempt to mitigate their impact when they become established. Pest risk maps are used by phytosanitary agencies to support risk analysis, pest surveillance, and emergency programs. One of the most important types of risk maps are those that estimate potential distribution based on climate suitability, which are usually created with bioclimatic models. Risk maps for non-indigenous species are generally created under two main constraints, limited time to produce the risk map product and sometimes gaps and uncertainties in the biological data needed to fit and validate the model. These constraints suggest that a suitable model for phytosanitary applications should be both generic and simple to use. One widely used approach is species distribution modeling, where a pest's potential distribution is inferred from a mathematical relationship between climate variables and the known distribution. One popular example is MaxEnt, which uses distribution data in combination with derived background observations (Phillips et al. 2006). However, MaxEnt and other species distribution models may not extrapolate reliably especially into novel climates (Elith and Leathwick 2009; Elith et al. 2012; Kriticos et al. 2014). An alternative is CLIMEX-Compare Locations (Sutherst and Maywald 1985), which uses literature and distribution data to fit the model parameters. CLIMEX is a simple process-based model that unlike a spatial distribution model, contains functions that explicitly define biological processes. CLIMEX is very widely used, partly because if there is insufficient literature data to parametrize the model it can be inferred from the distribution data itself. One disadvantage of CLIMEX is that it has many parameters and can be relatively labor intensive and subjective to fit, although an improved algorithm for auto-fitting could relieve some of these issues.

In addition to the need for potential distribution maps, other phytosanitary applications of weather or climate-based models include predictions of: i) the frequency

of years favorable to crop losses or epidemics (Pinkard et al. 2010b); ii) the timing of life stages to deploy surveys or treatments; iii) the duration of mitigation treatments designed to achieve control or eradication based on historical or forecast weather; and iv) the extent of crop damage or injury to specific hosts (Magarey et al. 2014; Magarey et al. 2007; Pardey et al. 2013; Pinkard et al. 2010b). These kinds of applications are also relevant to the management of indigenous pests. Process-based models have an advantage over the species distribution modeling approach in that they can be used to make these kinds of predictions. Though this does push the modeling process towards greater complexity as additional host and management factors are included. As evidence of this many process-based models developed for management of endemic pests, especially plant diseases can become quite complicated (Rossi et al. 2007). However, there is a trade-off. As models become more complex, they are increasingly difficult to adapt to a new species. As evidence of this there are hundreds of publications for CLIMEX a simple process model but for its sister product DYMEX a more complex generic model, there are far fewer published examples.

In order to address this problem, there is benefit in creating a simple generic model framework that is a compromise between ease of use and capabilities for additional phytosanitary and pest management applications. In this study, we introduce the Generic Pest Forecast System (GPFS), for simulating relative pest populations in space and time. The GPFS model presented in this study has the following components: i) Developmental rate estimated from cardinal temperatures (Sutherst et al. 1999); ii) Mortality from cold (Kaliyan et al. 2007), heat (Dentener et al. 1996), and soil moisture; iii) Population index based on developmental rate and mortality. Although these are the basic model components, the GPFS model also includes components for: iv) Infection and sporulation modules for plant pathogens; v) Pest and host growth stages based on degree days; and vi) Potential damage based on predicted pest population and host and pest growth stages (Magarey unpublished data), however these last three components will not be presented in this study. The GPFS model is designed to run within a pest information platform such as NAPPFAST (Magarey et al. 2014; Magarey et al. 2007) which would supply the required hourly weather inputs. The NAPPFAST system (used by the U.S. Department of Agriculture's Animal Plant Health Inspection Service between 2002 and 2014) included an interactive template to allow users to create simple degree day, disease infection, and flexible models from U.S. and global weather databases for phytosanitary applications. The GPFS model is a process-based model of the pest-host interaction and is not designed to simulate factors that may limit host distribution such as aridity. As a consequence these kinds of factors must be considered separately using additional climate layers inside a geographic information system.

The oriental fruit fly (OFF) (*Bactrocera dorsalis*) was chosen as a study pest to test the GPFS model because there is an extensive amount of literature data available for model development and validation. *B. dorsalis* lays eggs below the skin of the host fruit and develops from egg to adult in as little as 17 days but development can be substantially delayed under cooler conditions (Christenson and Foote 1960) The larvae feed on fruits and mature larvae drop to the ground and pupate in the soil. Adults typically

live for up to 3 months but may live longer in cooler conditions. Like other fruit flies, *B. dorsalis* requires favorable temperature and soil moisture conditions (Yang et al. 1994) and is one of the key pest groups of southeast Asia and Hawaii, causing damage to fruits and vegetables by larval feeding (Clarke et al. 2005). *B. dorsalis* is considered to be a species complex and *B. invadens* have recently been determined to be the same species (Schutze et al. 2014). *B. dorsalis* can be a major threat to agricultural crops because of extreme polyphagous behavior and is known to be highly invasive (Clarke et al. 2005). We conducted two types of evaluations of the GPFS predictions against *Bactrocera dorsalis* observations. The first was what we termed site-specific temporal validation to refer to the comparison of model predictions with observations from specific locations where data are collected at regular intervals over multiple years. This kind of validation is recommended by the developers of CLIMEX (Sutherst et al. 1999) and has been conducted in several CLIMEX studies (de Villiers et al. 2013; Legaspi and Legaspi 2010; Pinkard et al. 2010a). The second type of validation was a comparison of predicted suitability based on 10-years of weather data against the known distribution of *B. dorsalis*.

In summary, the objective of this study is to introduce the GPFS model and validate it for *B. dorsalis* using site-specific observations and distribution data. In addition, we wished to use the GPFS model to investigate the potential for establishment in the United States. No information collected on site was used to parameterize the model with the exception of food availability. In addition, no local weather data were used as input into the models to investigate the potential for gridded global hourly weather data to be used for historical pest predictions.

Materials and methods

Pest observations. Site-specific temporal pest observations were obtained from three studies in which adult oriental fruit fly were trapped (Han et al. 2011; Jayanthi and Verghese 2011; Vargas et al. 2010). The studies were selected because they contained multiple years of data and the observations could be obtained from the authors or read from the figures. No pesticides were believed to have been applied at these sites. The study sites include Hawaii in the United States, Wuhan in China, and Bangalore in India (Table 1). Hawaii has a subtropical climate with temperatures and humidity moderated by trade winds blowing oceanic air over the islands. Monitoring was conducted on Hawaii Island, HI, from September 2007 to March 2008 once every two weeks using methyl eugenol traps (Vargas et al. 2010). Wuhan has a humid subtropical climate with abundant rainfall and four distinct seasons. Observations were collected at an experimental farm in which various fruits and vegetable crops were growing. Adult male flies were sampled with methyl eugenol-baited traps from January to December in 2008 and 2009 at 10-day interval with replacement of lures every 20 days. The oriental fruit fly population data in Figure 1 of Han et al. (2011) were used for site-specific validation in China. Bangalore has a tropical

Table 1. Locations of case studies.

Reference	Location	Latitude	Longitude	Data period
Vargas et al. 2010	Hawaii Island, HI, USA	19.42942 (19°25'45.912")	-154.882 (-154°52'55.2")	2006–2008
Han et al. 2011	Wuhan, China	30.42915 (30°25'44.9394")	114.3639 (114°21'50.04")	2007–2009
Jayanthi and Verghese 2011	Bangalore, India	12.93686 (12°56'12.6954")	77.62111 (77°37'15.996")	1999–2002

savanna climate with distinct wet and dry seasons. Observations were collected in a guava orchard at the Indian Institute of Horticultural Research, Bangalore. The guava orchard was chosen to provide a food source for oriental fruit fly during the off-season for mango. Traps baited with methyl eugenol were used to monitor the insect population from June 2000 to June 2002. Population observation data from Figure 3 of Jayanthi and Verghese (2011) were used.

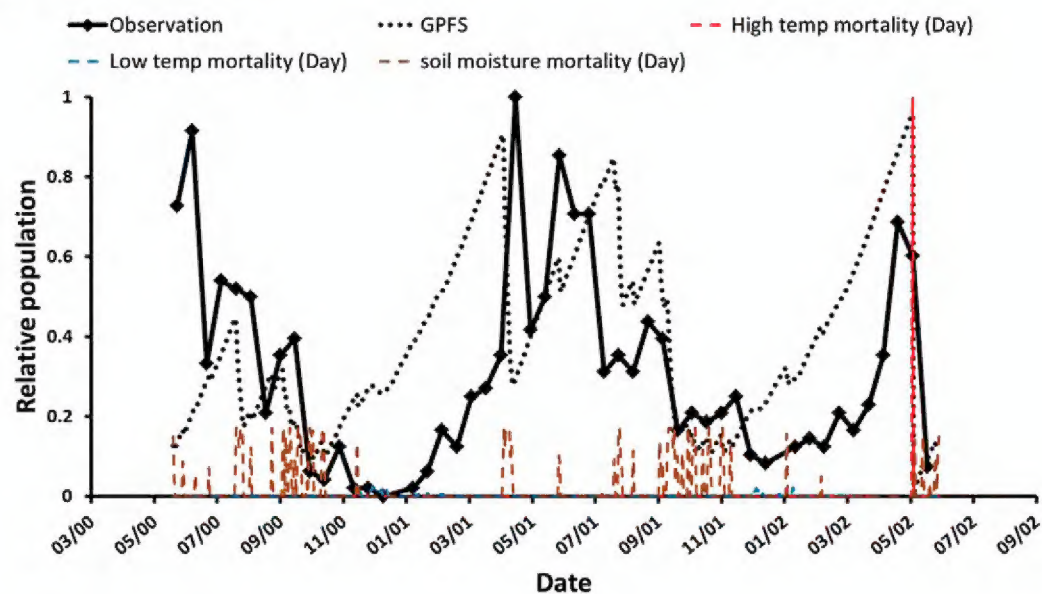
The Hawaii data were obtained from the authors whereas the observations for Bangalore and Wuhan were extracted directly from figures in the papers. Data extraction was conducted using a spreadsheet program (Excel 2010, Microsoft, Redmond, WA). The figure from the paper was scanned and copied into the spreadsheet and overlaid with a finer scale transparent grid-shaped graph with the same range of x- and y-axes to improve the ease of reading the data.

In addition to the site-specific observations, pest distribution data were also obtained from the literature (Suppl. material 1). *B. invadens* has recently be shown to be the same species as *B. dorsalis* (Schutze et al. 2014) so these observations were also included.

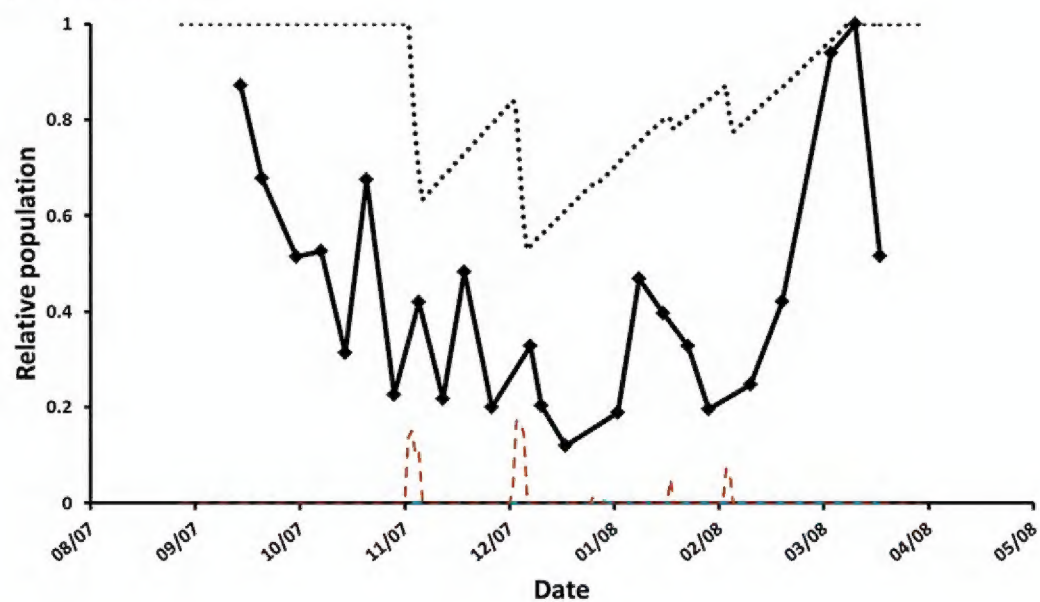
GPFS model. The GPFS model is designed as a simple generic tool for pest prediction for both arthropods and pathogens. The model is designed to run from hourly weather data inputs and to make predictions of the influence of weather on the relative pest population (population index) and phenological stages. For oriental fruit fly, GPFS only utilizes modules for development rate, high and cold temperature mortality, wet soil moisture mortality, and population index (Table 2). The first step in the GPFS model is to calculate the developmental rate in each hour. Next, the mortalities due to high temperature, cold temperature, and wet soil are calculated from hourly temperature and precipitation. The next step is to calculate the population index. The population index each hour is based on the sum of a development rate (scaled by the number of generations to reach maximum population) while removing the proportions of the population killed by high and low temperatures and by wet soil moisture. The population index is not sub-divided into individual life stages. Finally, the population index is adjusted by a simple function to account for lack of host availability.

Developmental rate. The hourly developmental rate (D) was estimated from four parameters: minimum temperature (T_{min}), lower optimum temperature (T_{opt1}), upper optimum temperature (T_{opt2}), and maximum temperature (T_{max}), describing the rate

A) Bangalore, India



B) Hawaii, U.S.



C) Wuhan, China

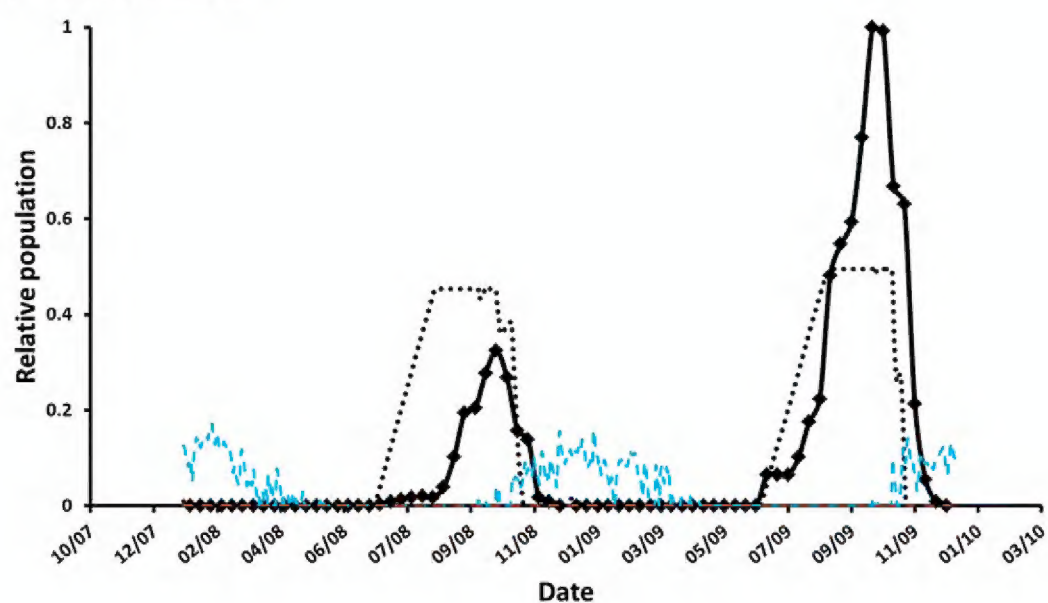


Figure 1. Comparison of observed (straight line with markers) and GPFS predicted (dashed line without markers) population of the adult oriental fruit fly, *Bactrocera dorsalis*, at three locations: **A** Bangalore, India **B** Hawaii, USA; and **C** Wuhan, China. Raw data of *B. dorsalis* field observations were converted to a population index (range: 0 to 1) to facilitate the comparisons

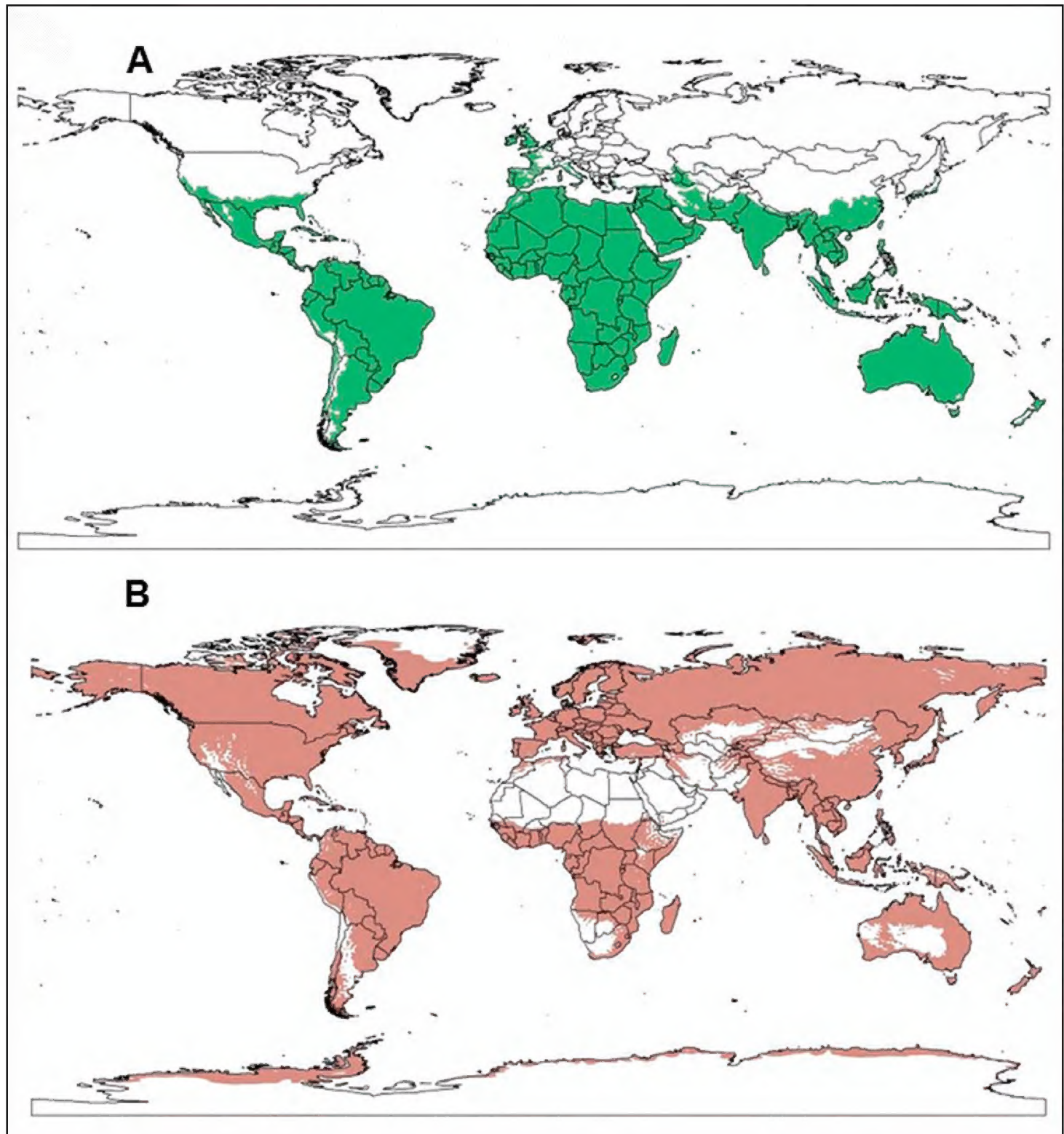


Figure 2. Cold and dry exclusions based on one or more occurrence of minimum temperatures of $-10\text{ }^{\circ}\text{C}$ (**A**) and annual precipitation less than 254 mm (**B**).

of development or population growth (Sutherst et al. 2007; Sutherst et al. 1999). We assumed that oriental fruit fly followed linear or constant development rates between these temperature thresholds. Developmental rates at each range obtained from the following equations.

$$\text{If } T < T_{\min}; \text{ or } T > T_{\max} \text{ then } D = 0 \quad (1A)$$

$$\text{If } T_{\min} < T < T_{\text{opt1}} \text{ then } D = (T - T_{\min}) / ((T_{\text{opt1}} - T_{\min}) * 24) \quad (1B)$$

$$\text{If } T_{\text{opt1}} < T < T_{\text{opt2}} \text{ then } D = (1/24) = 0.041677 \quad (1C)$$

$$\text{If } T > T_{\text{opt2}} \text{ and } < T_{\max} \text{ then } D = (T - T_{\text{opt2}}) / ((T_{\text{opt2}} - T_{\max}) * 24) \quad (1D)$$

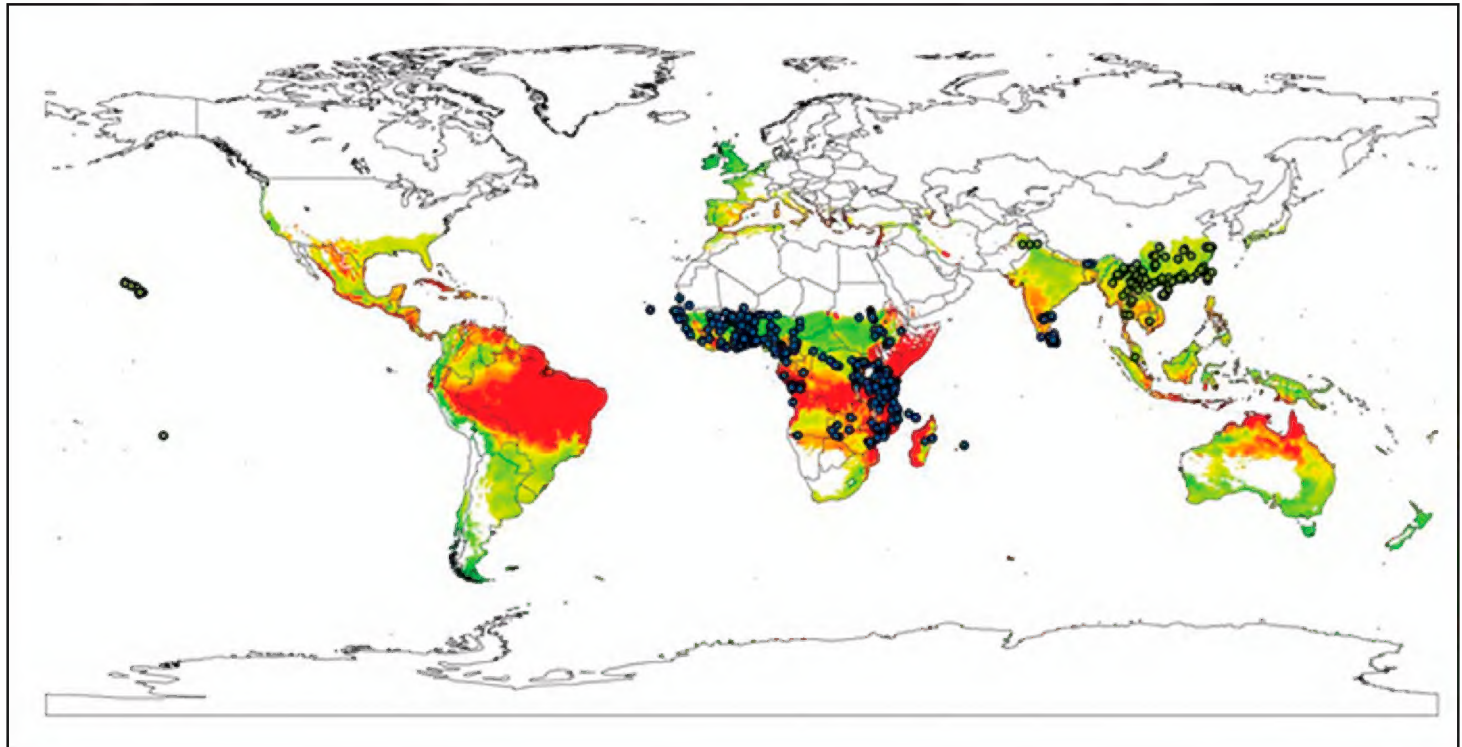


Figure 3. GPFS model prediction of potential global population index of oriental fruit fly, *Bactrocera dorsalis* (including observations of *B. invadens*), based on most recent 10 years (2003–2012) weather data from National Centers for Environmental Prediction – Climate Forecast System Reanalysis (NCEP-CFSR) at a 38 km resolution. The predictions are on a scale of 0-1 and do not account for the presence or absence of suitable hosts. The map is the highest population index of any month averaged over ten years with initial populations in each year being independent. The map also includes the cold and dry exclusions from Figure 2.

where T is the hourly air temperature, °C. Threshold temperatures of four developmental rate parameters were obtained from refereed papers: $T_{min} = 13.3$ °C (Christenson and Foote 1960), $T_{opt1} = 24$ °C (Vargas et al. 2000), $T_{opt2} = 34$ °C (Vargas et al. 1996; Yang et al. 1994), and $T_{max} = 41$ °C (Christenson and Foote 1960). When hourly temperature is between T_{opt1} and T_{opt2} , the rate of development reaches the maximum value of 1.0 per day or 0.041677 per hour.

Low temperature mortality. The hourly low temperature mortality was estimated from an exponential equation (Kaliyan et al. 2007). Although this equation was developed for Indian meal moth (*Plodia interpunctella*) it is based on the influence of temperature on the rate of chemical reaction and as such is a generic equation that can be adapted to other species. For temperatures less than T_{Mlt}

$$M_{lt} = [1 / (\text{EXP}(\beta_1 + \beta_2 / (T + 273.2)) + \beta_3 * \text{LN}(T + 273.2))] \quad (2)$$

where M_{lt} = the proportion of the population dying from low temperature mortality in an hour at temperature T in °C. To estimate parameters (β_1 , β_2 , and β_3), published data for lethal time (LT_{100}) of *B. dorsalis* were obtained (Burikam et al. 1992). However, due to the lack of information, additional data was included from *B. invadens* (now considered to be the same species) (Grout et al. 2011) and another species *B. tryoni* (De Lima et al. 2007; Heather et al. 1996; Jessup 1992; Jessup et al. 1993; Jessup et

Table 2. Parameters and abbreviation used in the GPFS model for oriental fruit fly.

Symbol	Parameter name	Value	Reference
D	Development rate		
T	Ambient temperature, °C		
T_{min}	Minimum temperature	13.3	Christenson and Foote (1960)
T_{opt1}	Low optimum temperature	24	Vargas et al. (1996)
T_{opt2}	High optimum temperature	34	Vargas et al. (1996)
T_{max}	Maximum temperature	41	Christenson and Foote (1960)
M_{lt}	Low temperature mortality		
T_{Mlt}	Threshold	13.3	Burikam et al. (1992), Grout et al. (2011), De Lima et al. (2007), Heather et al. (1996), Jessup (1992), Jessup et al. (1993, 1998)
β_1	Constant for M_{lt}	1101.7	
β_2	Coefficient for $1/(T+273.2)$ in M_{lt}	-49892	
β_3	Coefficient for $\text{LN}(T+273.2)$ in M_{lt}	-162.9	
M_{ht}	High temperature mortality		
T_{Mht}	Threshold	33	Armstrong et al. (2009), Jang et al. (1999), Xie et al. (2008)
β_4	Constant for M_{ht}	25.9595	
β_5	Coefficient of degree one for M_{ht}	-0.4959	
β_6	Coefficient of degree two for M_{ht}	0	
M_{wsm}	Wet soil moisture mortality		Eskafi and Fernandez (1990), Hou et al. (2006), Xie and Zhang (2007)
β_7	Constant for M_{wsm}	50.4	Eskafi and Fernandez (1990), Xie and Zhang (2007)
P	Proportion of population in soil inhabiting life stages for M_{wsm}	0.36	
$P_{(n)}$	Population index		
β_8	Reciprocal of number of hours required to reach maximum population for $P_{(n)}$	0.008	
β_9	Generations to reach maximum population under optimum conditions for $P_{(n)}$	4	
β_{10}	Hours to complete one generation under optimum conditions for $P_{(n)}$	30.4 d	
β_{11}	Extinction threshold	0, 0.0001	

al. 1998) were also used to estimate the parameters for low temperature mortality. The fit for low temperature mortality is shown in Suppl. material 2 and Suppl. material 5 (using Celsius units for ease of interpretation).

Puparia of *B. dorsalis* can survive freezing conditions and overwinter at low levels in China (Han et al. 2011); however, there were no experimental data available for parameter estimation. Instead, the threshold temperature for overwintering survival for creation of a cold exclusion mask was derived from comparisons of oriental fruit fly distribution records with extreme annual minimum temperatures. This analysis was

done outside of the GPFS model in a geographic information system and is described below. It also did not take into account the insulating influences of snow cover.

High temperature mortality. The hourly high temperature mortality is given by a polynomial equation with parameters β_4 , β_5 and β_6 fitted from observations of mortality under controlled conditions using an exponential function that has been shown to have utility for predicting heat mortality for light brown apple moth (*Epiphyas postvittana*) and Long tailed mealy bug (*Pseudococcus longispinus*) (Dentener et al. 1996). For temperatures greater than T_{Mht} ,

$$M_{ht} = (60/[EXP(\beta_4 + \beta_5 * T + \beta_6 * T^2)]) < 1 \quad (3)$$

where M_{ht} = the proportion of the population dying from high temperature mortality in an hour at temperature T . To estimate parameters (i.e., β_4 , β_5 , and β_6) of high temperature mortalities, LT_{100} of oriental fruit fly in response to extreme high temperatures, ranging from 43 to 50 °C, were obtained from published data (Armstrong et al. 2009; Jang et al. 1999; Xie et al. 2008). The studies were conducted in either a heating block system (HBS) or a forced-air chamber. LT_{100} data were fitted to the model of the denominator part in the equation (3). All parameters (β_{1-6}) were fitted using PROC NLIN in SAS software 9.3 (SAS Institute, Cary NC). The fit for high temperature mortality is shown in Suppl. material 2 and Suppl. material 4.

Wet soil moisture mortality. Excessive soil moisture i.e. flooding can reduce the populations of some fruit flies including *B. dorsalis* (Xie and Zhang 2007). The hourly wet soil moisture mortality M_{wsm} is given by a simple empirical relationship. If soil is flooded then,

$$M_{wsm} = P*(1/\beta_7); \text{ else } M_{wsm} = 0 \quad (4)$$

where M_{wsm} = the proportion of the total population dying from soil moisture mortality in an hour, and P = the proportion of the population in soil-inhabiting life stages. This was assumed to be a constant 0.36 based on the pupal proportion of total degree days. The parameter β_7 is the number of hours that the life stage will survive in flooded soil. In the absence of soil moisture data, the soil was defined as flooded if more than 10 mm of rain had fallen in the previous 24 hours. Parameter (β_7) for soil moisture was also estimated by calculating LD_{100} in flooded soil from published data with *B. dorsalis* (Xie and Zhang 2007) and *Ceratitis capitata* (Wiedemann), Mediterranean fruit fly (Eskafi and Fernandez 1990). Mortality (LD_{90}) was 2.49 d for third instar *B. dorsalis* larvae in flooded conditions at 25 °C (Xie and Zhang 2007); however, no information for the pupae was available, so it was assumed that pupae would respond similarly to flooding like larvae. *Bactrocera dorsalis* pupae did not survive at soil moisture greater than 80% (Hou et al. 2006), although the survival time in saturated soils is unknown. A study with Mediterranean fruit fly showed that survival under similar conditions were 4 and 3 days for larvae and pupae, respectively (Eskafi and Fernandez 1990). We used this data to estimate approximate LD_{100} of *B. dorsalis* pupae under saturated con-

ditions and calculated LD_{100} as $(2.49/0.9) \cdot (3/4) = 2.1$ day. Thus, $(\beta_7) = 50$ hours. Wet soil moisture mortality does not account for increased pupal survival that occurred with increased soil humidity (Vargas et al. 1987), but only mortality was associated with flooding. Since this version of the GPFS model did not calculate the proportion of the population in each pest stage, we assumed a fixed proportion (0.36 based on stage duration) of the population was present in the pupal stage at any point in time.

Population index. The population index is a measure of relative population as influenced by weather conditions and is a function of the developmental rate, the mortality rates and the population index in the current hour (n).

$$P_{(n)} = [P_{(n-1)} + (\beta_8 * D)] * (1 - M_{lt}) * (1 - M_{ht}) * (1 - M_{wsm}) \leq 1 \quad (5)$$

where $P_{(n)}$ = the population index (0-1) at hour n . The parameter β_8 is the reciprocal of the number of hours required to reach the maximum value of the population index. It can be estimated from

$$\beta_8 = 1/(\beta_9 * \beta_{10}) \quad (6)$$

where β_9 = generations to reach maximum population index under optimal conditions and β_{10} = days to complete 1 generation under optimal temperature conditions. The variable β_9 was arbitrarily chosen based on the assumption that a minimum of four generations would be required to reach the maximum value of the population index. The model also includes an extinction value, β_{11} , which if the final population falls below due to mortality, the population would remain at 0 even when favorable conditions returned.

Host fruit availability. The period when host fruits are available is an important factor in determining *B. dorsalis* population increase (Chen and Ye 2007). Monthly status of fruit availability (1: food, 0: no food) was implemented into the GPFS model to adjust population size based on fruit tree phenology. If food was not available for a given month, then the population was allowed to decrease during unsuitable periods, but not increase. In Wuhan, China, we set food as available from July to December (Han et al. 2011). Peaches were the only fruit available during May and June, but no oriental fruit fly larvae were found in peaches, so these months were not included as having an available food source. No food limitation was applied to Bangalore, India where the combination of mango and guava likely provides a near year round source of food and Hawaii, USA, where there are primary hosts, such as strawberry guava and common guava, as well as other fruits and vegetables (Cornelius et al. 2000; Vargas et al. 1983; Vargas et al. 1990).

Pupal cold mortality. There is evidence to suggest that pupae survive lower cold temperatures than larvae. In Wuhan, pupae were shown to successfully overwinter although survival was dependent upon the time of year when pupae were placed in the soil (Han et al. 2011). In addition, the authors also found occasional pupae in the field. Although, the authors did not definitively demonstrate these occasional pupae

may give rise to overwintering populations, there is reason to err on the side of caution when constructing models of potential distribution, especially those for phytosanitary applications. Since there were insufficient data to parameterize the cold mortality of pupae, this was simulated outside of the GPFS model framework by comparing the extreme annual minimum temperature with the oriental fruit fly distribution maps, assuming this would represent the range of OFF cold survival or seasonal migration. A probability map showing the frequency of -10°C or less each year was used as the exclusion layer and to mask out areas where oriental fruit fly would not overwinter. The Extract by Mask function in ArcGIS was used to create a global map showing likely maximum population and areas where oriental fruit fly would not survive. In NAPPPFAST (Magarey et al. 2014), a probability map for one or more occurrences of minimum temperatures of -10°C or less from the period January 1 to December 31 was made using ten years of hourly CFSR weather data (2003–2012). A frequency of 2 or more years in 10 was considered to be likely to eliminate oriental fruit fly populations (since populations might recover by re-introduction if killing periods occurred only once in every ten years). The grids were imported into ArcGIS and compared with oriental fruit fly distribution records. The grid included all the distribution records in the zone with values of 1 year or less. A mask of unsuitable area (2 years or more) was created and then multiplied by the maximum population grid to create the final global oriental fruit fly map. Large lakes including the Caspian Sea were excluded from the final product.

Dry exclusion. In addition to cold exclusion, a dry exclusion map was generated to mask dry/desert areas from the global distribution map using the ArcGIS. If annual precipitation was less than 254 mm, then the areas were defined as arid and unsuitable for oriental fruit fly habitats. This limit is commonly defined as limit for aridity (Maliva and Missimer 2012) and represents areas on the map that are likely not to have suitable habitat for OFF.

Model runs. The hourly weather data used for the model comparison and the creation of maps were produced by the National Center of Environmental Prediction Climate Forecast System Reanalysis (NCEP-CFSR) at a 38 km spatial resolution (Saha et al. 2010). A commercial weather data company (ZedX Inc., Bellefonte, PA) provided hourly weather data for specific years and locations from the gridded CFSR data sets created by NCEP. The input variables for the GPFS model were hourly air temperature, and precipitation. For the site-specific comparisons, the model was run in MS Excel. For spatial comparisons, a version of the model was coded into the GNU Compiler Collection (Free Software Foundation, Inc. <http://gcc.gnu.org>). Risk maps were created by running the GPFS model from January 1 to December 31 using ten years of CFSR weather data. To investigate the maximum climate suitability, the model was run using the same initial population index value in each year and populations did not carry over from one year to another. The model was run with the extinction value (β_{11}), set at 0. A map of final population in the middle of each month for each year was created and the average was calculated for each month over the ten year period. Each of these monthly maps was imported into a Geographic Information System for fur-

ther processes (ArcGIS, ESRI, Redlands, CA) including incorporation of cold and dry masks. Since populations may peak at different times of the year depending on climate, a summary grid reporting the maximum population was created. For a risk map specific to the conterminous United States, Real-Time Mesoscale Analysis (RTMA) hourly data at 5 km spatial resolution (Benjamin et al. 2007) were used instead of CFSR. The main component of RTMA is the Gridpoint Statistical Interpolation (GSI) system derived from over 14,000 weather station observations, radar and satellite observations in NCEP (De Pondeca et al. 2011). Observations are ingested each hour and a 1-hour forecast is used as background layer to supplement observations for the next iteration. To look at the potential for pest establishment, the GPFS model was run from 2007 to 2012 with an initial population index in January 2007. In each subsequent year, the initial population index was the final population index in the previous year. The final map represented the maximum population index of any month in the final year (2012) and included the cold mask.

Site-specific temporal comparisons. Several accuracy measurements were calculated to determine how well GPFS predictions fit the observed population changes at the study locations. To facilitate comparisons with the predicted populations, the trap catch data were scaled between 0 and 1. The scaled values were calculated by dividing each observation by the 99th percentile of data from all years at each location. The statistical tests included mean error (ME), mean absolute error (MAE), mean square error (MSE), square root of MSE ($RMSE$), and index of agreement (d). Definitions and interpretations of these indices are well described in Legates and McCabe (1999) and Moriasi et al. (2007) and the equations for computing these indices were adopted or modified from these literatures. ME , MAE , MSE , and $RMSE$ are indices for error that describe the difference between model prediction and observations in the units (or squared units) of the variable. Equations for computing these indices are expressed as:

$$ME = \left[\frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{pred})}{n} \right] \quad (7)$$

$$MAE = \left[\frac{\sum_{i=1}^n |Y_i^{obs} - Y_i^{pred}|}{n} \right] \quad (8)$$

$$MSE = \left[\frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{pred})^2}{n} \right] \quad (9)$$

The index of agreement (d) measures the degree of how observed values are accurately estimated by the simulation. The index of agreement is different from a measure of correlation or association in that it measures the degree of error-free of the model's predictions (Legates and McCabe 1999; Willmott 1981). Like correlation coefficients,

it ranges between 0 and 1, where a value of one indicates a perfect match between observed and simulated variables while zero value suggests a complete disagreement. The index of agreement (d) is expressed as:

$$d = 1 - \left[\frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{pred})^2}{\sum_{i=1}^n (|Y_i^{pred} - Y^{mean}| + |Y_i^{obs} - Y^{mean}|)^2} \right] \quad (10)$$

Spatial distribution comparisons. Model accuracy measures, modeled prevalence and sensitivity, were estimated using the final GPFS risk map. Raster cell values were extracted using ArcGIS. No data values (i.e., -9999), mainly assigned to oceans, were excluded from the analysis. The modeled prevalence is the proportion of raster cells classified as suitable. To estimate prevalence, the least presence threshold was used to classify raster cells as suitable or unsuitable. In species distribution modeling, the lowest presence threshold (LPT) is commonly defined as the predicted value of lowest training observation (Webber et al. 2011). We used the predicted value of 1st percentile in the training observations so the threshold is not influenced by an individual outlier. The modeled prevalence was estimated by dividing the number of cells with values greater and equal to LPT value by total number of raster cells. The model sensitivity is the proportion of test locations falling in suitable raster cells.

Results

Site-specific temporal validations. At the China and India sites, the GPFS model populations went extinct due to cold and/or heat mortality with the extinction threshold set at 0.0001. Since the parameters for pupal cold mortality were unknown, the cold mortality threshold was not able to be estimated. Instead the model was run with an extinction threshold set to 0. For high temperature mortality, a correction was made to the threshold, T_{Mbt} to 39 °C from the literature value of 33 °C, which improved accuracy and allowed populations to persist in Bangalore. This may indicate that OFF individuals can move to find more favorable microclimates and thus avoid the highest temperatures. All other parameters remained the same from their literature values. The highest index of agreement was 0.85 at Wuhan. Regardless of modeling systems, the mean errors between the scaled adult populations and the predictions were smallest at Wuhan, followed by Bangalore and Hawaii.

The GPFS population predictions at Bangalore, India matched relatively well with the observed population (Figure 1A). The GPFS model correctly simulated the major peak populations in 2001 (April 2) and 2002 (April 22). Although the first peak in 2000 (June 19) was not correctly simulated, it may be due to increased food availability associated with mango harvest. In 2000, the GPFS modelled a peak in April 21 (Data not shown in figure). The main drivers of population decline in the GPFS model run at Bangalore were cold mortality during the cooler months and soil moisture mortality during the wetter summer months (including the monsoonal

months, typically August to September). Heat stress mortality did not seem to be a population limiting factor.

For Hawaii, USA, the GPFS model simulated population dynamics relatively poorly compared to the other two locations (Figure 1B). There are several reasons for this. The first is that temperatures never reach extremes that cause steep population declines. Second, even by Hawaii standards this area has high populations due to the presence of many hosts throughout the trapping area. Methyl-eugenol traps, because they draw from a very broad area, can be difficult to interpret under these conditions. Consequently, distinct peaks in trap captures did not occur like the other sites (Figure 1). The GPFS model predicted a population decline during the wettest part of the year but underestimated the magnitude of the decline. The reduced climate suitability was driven mostly by excessive soil moisture. The model also had food as available all year round in Hawaii, although it is possible that there is less food available in the winter months. Another factor is that there are more uncertainties in weather data in Hawaii due to few stations and topographical influences.

The GPFS predictions matched the observed oriental fruit fly population comparatively well in Wuhan (Figure 1C) compared to the other two study locations (i.e., Bangalore and Hawaii). The GPFS model overestimated the mean population in 2008, but underestimated it in 2009. In 2008, the simulated population increased faster and had a higher peak compared to the observations. This may be due to the fact that surviving overwintering population levels are very low. The model concordance might be improved by a finer temporal representation of food availability. For example, the first host that supported oriental fruit fly oviposition, pears, did not ripen until mid-July, while the model allowed food to be available throughout July. This consideration is balanced by the fact that some OFF oviposition may occur before this period on unripe fruit. OFF prefer ripening fruit and survival is lower on unripe fruits (Chiu and Chen 1987). Declines in the GPFS simulated population in Wuhan were driven entirely by cold temperatures since heat mortality was not a factor.

Spatial distribution validations. The cold and dry exclusions eliminated large portions of Northern Europe, Asia and America and desert regions of Africa and Asia (Figure 2). The cold exclusion was based on the frequency of -10°C or less each year and the dry exclusion was based on areas receiving less than 254 mm. The GPFS model was run globally at a 38 km grid resolution with extinction value β_{II} set at 0 (Figure 3) and 0.0001 and with January 1 in the Northern Hemisphere and July 1 in the Southern Hemisphere (data not shown) start dates. For the locations where oriental fruit fly has been observed in Asia (Figure 4B), the mean and median of predicted maximum population were 0.404 and 0.338, respectively, with 25th and 75th percentiles of 0.197 and 0.577. With the threshold set at 0, the model predicted low levels of oriental fruit fly population surviving in Northern China in the middle of December, well beyond the range of the observed oriental fruit fly distribution. With the threshold set at 0.0001 the model greatly under predicted the range of oriental fruit fly in China. However, as discussed earlier, pupal mortality from cold temperatures may be much lower than larval mortality based on observations of pupal survival (Han et al. 2011).

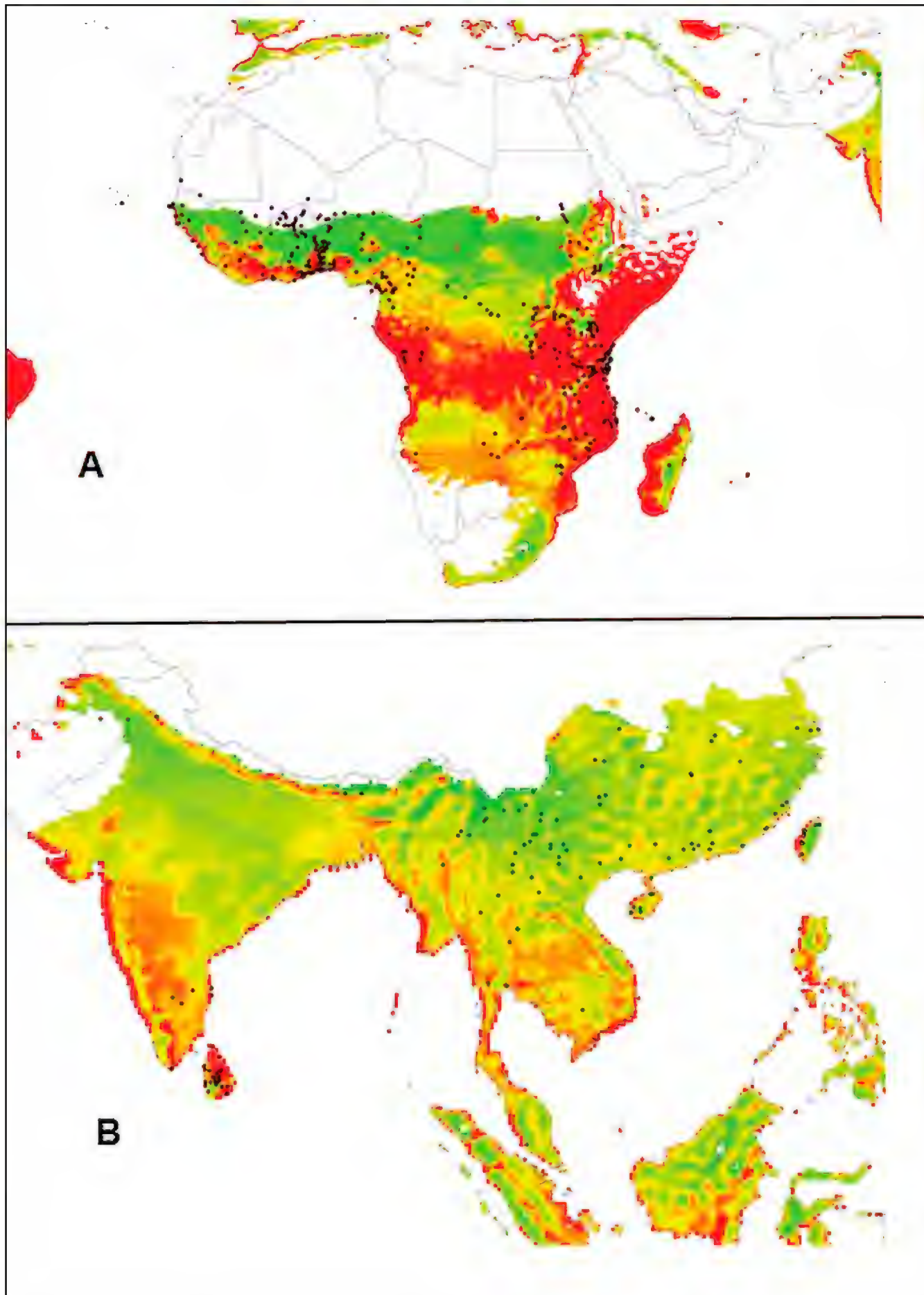


Figure 4. GPFS model prediction for potential population index of oriental fruit fly, *Bactrocera dorsalis* (including observations of *Bactrocera invadens*) in A) Africa and B) Asia based on most recent 10 years (2003–2012) weather data from National Centers for Environmental Prediction – Climate Forecast System Reanalysis (NCEP-CFSR) at a 38 km resolution. Locations where *B. dorsalis* or *B. invadens* has been observed in the literature are shown as black dots. The predictions are on a scale of 0–1 and do not account for the presence or absence of suitable hosts. The map is the highest population index of any month averaged over ten years with initial populations in each year being independent. The maps also include the cold and dry exclusions from Figure 2.

Since pupal survival parameters were not known the global suitability map included the cold and dry exclusions calculated independently and used to mask out unsuitable areas (Figure 2). Since most oriental fruit fly observations were in China, we show the map for this region (Figure 4B). The maps suggested that the Least presence threshold (LPT) based on the GPFS model output was 0.055, which was surprisingly low, although species distribution records are often collected for rare species (Bradley 2013). The model sensitivity was 0.99, i.e., over 99 % of observations (159 out of 161) were found within areas modelled as being climatically suitable (Figures 3 & 4). In Asia, the GPFS model predicted most oriental fruit fly occurrences in suitable regions except one in the northeastern Pakistan (Figure 4B). In Africa, the fit of the model against *B. invadens* observations was good with the exception of a number in the Sahel of Africa that were excluded from the suitable range because they were considered too dry. Globally, the modeled prevalence was 0.21, but without the cold and dry exclusion it was 0.95. The analysis assumed that host plants are available globally. The global climate suitability map indicates that oriental fruit fly may survive in most south parts of Africa and America, whereas only in limited areas of Europe, North America and North Africa are suitable due to cold or dry weather conditions (Figure 3). In parts of Europe, southern parts of South America and southern Australia, the climate for OFF is marginally favorable. Populations may not well survive in these areas during unfavorable years. Running the model in the establishment mode i.e. with population index change after five or more years may answer this question. However, it was not our objective to evaluate this question except for the United States where this was completed with RTMA data. The climate suitability map for the U.S. using 5 km resolution RTMA data suggests that potential fruit fly distribution may be limited to southern and western coastal areas in the United States (Figure 5).

Discussion

In this study, we introduced a new pest prediction model, the Generic Pest Forecast System (GPFS) and validated it against site-specific observations and spatial distribution. Importantly, no site-specific information was used to parameterize the model, with the exception of host fruit availability used in both models. In addition, no local weather data were used as inputs into the models to investigate the potential for a gridded global weather database to be used for historical pest predictions.

The goal of the team developing the GPFS model was to create a simple weather-based pest model that would have application for predicting potential distribution. In addition, the model was conceived also to have application to other risk based questions such as time of pest emergence and potential impacts in managed crop systems for both indigenous and non-indigenous pests. This additional information may enable decision makers to better understand the consequences of a newly established pest. One of the precedents for the GPFS model is weather-based pest forecast models which are routinely used in pest forecasting (Magarey et al. 2001). However, these

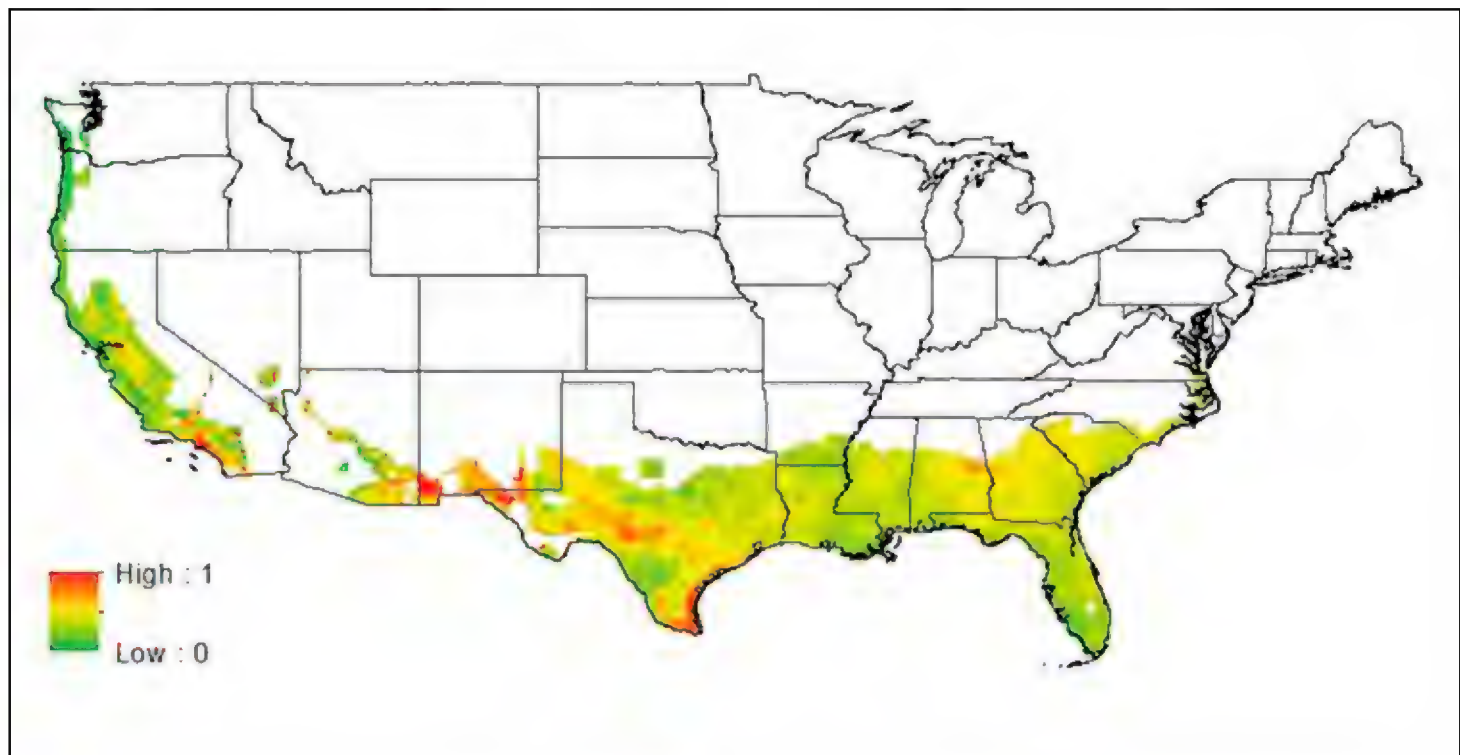


Figure 5. GPFS predictions of potential population index of oriental fruit fly, *Bactrocera dorsalis*, in the United States based on A. Real-Time Mesoscale Analysis weather data (2007–2012) with a 5 km resolution. The predictions are on a scale of 0–1 and do not account for the presence or absence of suitable hosts. The map is the final population index in December 2012 that resulted from the simulation of population change beginning with an initial population in January 2007. The map also includes the cold exclusion from Figure 2.

kinds of models do not typically include mortality factors since the pest is endemic and overwintering-survival is not usually in question. The GPFS also shares some similarities with CLIMEX-CL (Compare Locations), a widely used pest modeling software. CLIMEX is a relatively easy way to create a climate suitability map for a species since the generic nature of the growth and stress parameters enable the model to simulate the growth and survival of a species irrespective of the exact biological processes. This is probably one of the key reasons for the widespread use and adoption of CLIMEX-CL. In the GPFS, Eq. 1 is used to calculate developmental rate and in CLIMEX-CL, the same equation is also used in combination with a degree day formulation to calculate the temperature growth index (Sutherst et al. 2007). CLIMEX and GPFS both use stresses or mortality due to cold and heat but with different equations. The GPFS is designed to run from hourly weather inputs whereas CLIMEX runs from weekly indices usually estimated monthly climate averages. In GPFS, the population index accumulates each hour and is reduced by mortality factors. In CLIMEX, the Ecoclimatic Index (EI) value is calculated each week as the product of stress and growth indexes and then the EI is accumulated over the year. CLIMEX also includes additional types of stresses, and indices such as diapause and irrigation not considered in the GPFS. To compensate, the GPFS model may need to be used in combination with other GIS data sets when comparing pest distributions, such as those defining aridity.

The GPFS model appears to have a number of useful features. It has a relatively simple formulation, few parameters and can be used to investigate population changes dur-

ing the seasons as well as produce overall suitability maps. The simplicity of the model makes it easy to adapt to a new pest, assuming these parameters are available or can be determined from a closely related pest (the latter which would increase the uncertainties associated with the simulation). Adaptability of a model to a new species is important since phytosanitary agencies must often develop risk assessments at short notice.

Although GPFS is a generic model, it is still not as flexible as the generic pest simulation model in the CLIMEX family, DYMEX (Maywald 2007), in building a model for a targeted species or adopting an already existing model to another new species. However, an experienced DYMEX user is required to use the model the development portions of DYMEX (DYMEX model builder), although the task may be simplified if a model exists for a related species. The GPFS represents a compromise in terms of model complexity. For example, to calculate the developmental rates, simple linear slopes at each temperature threshold range were used. Many other modeling tools use unique pest specific equations to represent developmental rates (Gutierrez et al. 2010; Maywald 2007). However, linear estimation of developmental rate seems appropriate when there is scarce information on pest development data, especially for exotic species. The GPFS model presented in this study has the potential to be improved with additional modules for diapause and pesticide timing.

The ease of parameterization is another key consideration for the use of a pest model. One limitation for the modeling of exotic species is the lack of published experimental data to parameterize the model. A useful feature of CLIMEX-CL is that parameter values can be estimated from the experimental literature, from the distribution data or as recommended using both in combination. A disadvantage of CLIMEX-CL is the weekly model time step, which can make the parameterization process more difficult since some experimental data such as mortality may have observations made at an hourly time step. The GPFS requires biological parameters, including information on development rate and mortalities due to heat, cold, and wet soil moisture. Consequently, it would not be possible to make a GPFS model for some pest species at present. One option is to parameterize the model from a related species, which increases the level of uncertainty associated with the modeling process. One possible solution is to use the distribution data to fit the parameters using an automated process. Such a method has been employed to fit a mechanistic model to predict forest species distributions (Higgins, Mullen et al. 2011).

Accuracy is another key consideration for evaluating models. The results from this study need to be kept in context with the limited amount of site-specific information (including weather data) that was used to inform the model. Information on other factors such as pest migration, food quantity, species competition and/or existing natural enemies, and management factors were not included. A more sophisticated simulation model may be able to have superior results using this kind of site-specific variables. When these kinds of site-specific parameters are included in a simulation model it is prudent to check the model is portable to other sites (Yonow et al. 2004).

The GPFS model might also be improved by adding modules to account for the population of individual pest stages. This would allow it to be used for other applica-

tions for example predicting the timing of a particular phenological stage for scheduling trapping or scouting applications. Another module that could be improved is the food supply model, which was highly simplistic and also relied on a monthly time-step. In some cases the phenological susceptibility of a host may be more precisely estimated from an observed biofix and a degree day model. In phytosanitary risk analysis, there is sometimes a need to simulate potential impacts and spread of pest in order to assess mitigation or response options (Waage et al. 2005). There is potential to improve the GPFS model to account also for factors such as host development, pesticide treatments, host resistance and natural enemies; factors that have been included in other pest forecast, simulation models or expert systems (Fitt et al. 1995; Gutierrez et al. 2010; Travis et al. 1992). Estimates of spread and impacts based upon a simulation model might improve upon those inferred from a pest's native range where climate, host and management conditions might be quite different. Examples of climate based simulation models being used in these kind of impacts assessments suggest there is benefit in adding this capability (Kriticos et al. 2003; Pinkard et al. 2010c).

One limitation with process models that require hourly weather inputs has been the lack of reliable, consistent, and dense global historical weather station data for use in validating models against published historical data of pest populations. This situation changed in 2010 when the National Center of Environmental Prediction (NCEP) introduced the Climate Forecast System Reanalysis (CFSR) (Saha et al. 2010). The CFSR database superseded the NCEP R2, which had a crude spatial resolution of 200 km, and included a more simplified description of atmospheric processes. The CFSR database provides a relatively high resolution (38 km) source of high quality gridded data from 1982 until the present at a 1-hourly time step. It is believed to be the most consistent and reliable source of global gridded historical data. The CFSR also has a much more detailed description of the atmosphere and atmospheric processes and improved accuracy, especially in the Southern Hemisphere. In a comparison with 28 weather stations in horticultural production areas, CFSR data had a mean relative and mean absolute errors of $-0.3^{\circ}\text{C}/-4.3\%$ and $2.4^{\circ}\text{C}/12.8\%$ respectively for air temperature and relative humidity (Magarey, unpublished data). The reliability of CFSR data may also be dependent on other factors such as topographical complexity and the density of weather stations used as inputs into the reanalysis. It would have been informative to have compared CFSR weather data with those from local weather stations, but this was outside of the scope of the project. We did not do this in this study because of the complexity and some degree of spatial uncertainty around the exact locations of the field trials. CFSR data show promising performance for hydrological studies including stream flow and crop yield modeling (Dile and Srinivasan 2014; Fuka et al. 2013). The challenge for pest forecasting is potentially far greater due to the need for predictions data to be generated on finer spatial and temporal scales than those used for hydrology. To the best of our knowledge, this study is one of the first to use CFSR data for pest forecasting. The CFSR data set allows pest modelers not only to create high resolution global maps from hourly data but also to create historical predictions for which field observations at specific locations over multiple years have been published in the literature.

Utility of site-specific temporal validation

Site-specific temporal observations can play a useful role for model validation by providing an additional level of confidence that the model is providing realistic results. These types of validations may not be possible or can be difficult for some species due to pest and host phenology. That is that the number of individuals caught in a trap is a function of pest or host phenology and not just pest abundance. In this version of the GPFS model, we ignored pest phenology although it is incorporated into a newer version (Magarey unpublished data). However, including pest stages into the model greatly adds to complexity including the required parameters. In addition, for a pest with many generations and overlapping stages such as OFF it may not contribute much additional information. For example, including a stage specific model into the GPFS did not improve prediction accuracy for light brown apple moth (*Epiphyas postvittana*) (Hong and Magarey, unpublished data). The prediction of pest population index at specific sites can be revealing. For example, it might help identify sites where a pest is not expected to overwinter even if the summer climate is suitable. This is important since some distribution (spatial) records may be unreliable or the record may have corresponded with a rare or ephemeral observation of the species (Macfadyen and Kriticos 2012).

One important factor for site-specific validation is choosing representative locations. For example, predicted population indexes of *B. dorsalis* oriental fruit fly population in Wuhan, China were relatively superior compared to the two other locations (Figures 1C and Table 3). One likely cause of this is that cold winter conditions in Wuhan (absent at the other two sites) help synchronize population development. This also points out a limitation of site-specific validation in that it may not always be possible to obtain pest observations from locations that represent the potential range of suitable climates for an invasive pest. In addition, sites that have moderate climates such as Hawaii may not provide ideal validation sites for models to be deployed in continental climates because the populations are not primarily limited by climatic conditions. Another limitation of the site-specific validation is that the contribution of migration or human-mediated transport is not always clear. For example in Wuhan, populations potentially could overwinter at low levels (Han et al. 2011); however the contribution of fruit transported by humans may be a more important factor in the population cycle. Another limitation of site-specific validations is that observations lag predictions based on weather data. It might be possible to calculate a lag factor based on developmental time, however this could be complex depending on the biology of the pest. Since the fit was relatively robust without such a correction, we did not attempt to calculate a lag time. In addition, another limitation is population increases caused by abundance of food during periods such as harvest, Since the GPFS model is simple, it did not consider food quantity only the timing of food availability.

Utility of spatial distribution validation. Since site-specific temporal validation often includes relatively few locations, spatial validation is critical. Spatial validation of the GPFS showed the utility of the GPFS model but also the need for the cold exclu-

Table 3. Model accuracy measures between observation and predictions of GPFS.

	Bangalore, India	Hawaii, USA	Wuhan, China
Observation			
N	50	24	73
Mean	0.32	0.44	0.12
Standard deviation	1.75	1.18	1.77
GPFS			
Host/food availability	Jan–Dec	Jan–Dec	Jul–Dec
Prediction mean	0.40	0.85	0.13
Mean absolute error (MAE) (≥ 0)	0.23	0.41	0.08
Mean error (ME) (-Inf to +Inf)	-0.08	-0.41	-0.02
Mean square error (MSE) (≥ 0)	0.08	0.20	0.02
Index of agreement (d)	0.58	0.50	0.85

sion layer to prevent over-prediction in higher latitudes due to the lack of experimental data to parameterize the low temperature mortality. This shows that the combination of both complex (i.e., GPFS) and simple (i.e., exclusion) modeling techniques may be useful for defining the non-suitable areas (Figures 2 & 3). A recent study has shown that *B. dorsalis* is genetically indistinguishable from *B. invadens* (Jose et al. 2013). *B. invadens* is distributed from Senegal through the southern Congo with most occurrences in Benin and Cameroon along the equator in West Africa (Goergen et al. 2011). The GPFS model based on *B. dorsalis* parameters predicted these areas as suitable. The GPFS model also predicted that Angola and Namibia would also be suitable, although these countries likely have lower host densities. One note of caution is that populations of *B. dorsalis* from different localities may differ in their temperature and moisture requirements, so any predictions for an invading population should be treated with caution. This is especially the case since *B. dorsalis* is a species complex.

The GPFS predicted population can also be compared to a description of *B. dorsalis* populations found in an environmental chamber study in which temperature and humidity were maintained at levels representative of six U.S. cities (Flitters and Messenger 1953). This comparison of both the environmental chamber and the GPFS simulation showed that Fort Pierce, Florida and Oceanside, California were the most suitable for the oriental fruit fly, followed by Riverside, California. The chamber representing Fresno, CA, had low populations during the summer periods, while GPFS predictions were 0.13 and 0.17 in July 15 and Aug 15, respectively. In the chamber study and the GPFS model simulations using RTMA data (Figure 5), Vincennes (Indiana) populations could multiply in the spring and summer, but died out in the winter. In the Charleston (South Carolina) chamber, observed fly populations declined to zero during winter. The GPFS model suggests that oriental fruit fly could survive in Charleston, with the possible explanation that the model did not run through the end of winter, but stopped at December 31, prior to calculating averages for the ten year period. An alternative way to run the GPFS model is to begin with a founder popula-

tion and then run the model for a ten year period. The disadvantage of this approach is that it may underestimate potential distribution as population extinctions might be caused by one or more extreme events. For a pest that may be frequently introduced, the average of a ten year prediction might be the most suitable. Alternatively, the analyst may choose to select individual years based on the occurrence of extreme weather conditions that might limit pest establishment.

Ideally, host abundance and phenology are also required for prediction of the population index (or relative abundance), but this information is very rarely available, especially on global scales. We investigated the use of global crop maps (<http://capra.eppo.org/maps.php>) (Monfreda et al. 2008); however, we chose not to use them due to the poor reliability in certain areas and the large host range of oriental fruit fly, which added to the complexity and uncertainties of making the maps. As a consequence of the lack of host availability data, the GPFS tends to over-predict fruit fly populations in some areas. Another cause of over prediction is the lack of host plants; for example, in much of semi-arid Australia host plants are likely to not be abundant. However, for a local scale map such as a pest detection map for California or Florida, it would be important to include maps of cropland or host distribution, rather than using the exclusion based on precipitation to define host areas.

Conclusions

The GPFS model was introduced as a simple weather-based model for predicting potential distribution. The model was shown to be able to simulate relative pest populations in some locations, which could have potential to estimate potential impacts of a pest when combined with other biological and management variables. The model requires literature data to estimate model parameters and as such will not be usable for all species unless alternative methods of parameterization are added in improved versions of the GPFS model. This study also shows the potential for improving pest risk models by conducting spatial and site-specific temporal validations against published observations. Although these kinds of temporal validations will not be possible for every species, they can provide insight into the spatial domain by suggesting why a species might not persist or provide an indication of the risk. It can also be helpful for calibrating model parameter values. Importantly, the arrival of high quality global gridded historical weather databases can make site-specific temporal validations from published observations easier for hourly weather-based models. The downloading and archival of gridded data sets are a large undertaking requiring considerable resources. However, smaller organizations have the opportunity to purchase hourly data sets from commercial weather providers for specific sites of interest. Ultimately as computer power improves these costs will decrease. Additionally, the model could be run in real time to support surveillance activities given the concern about low level or cryptic invasions escaping pest detection programs (Papadopoulos et al. 2013). We suggest that models be validated both spatially and temporally, when possible, in order to increase

the confidence in their results. Caution is needed since in some locations a pest may not be coupled to climatic conditions and the population may be driven by other temporal factors, e.g., food availability. Spatial validation provides confidence that the model is working correctly in a range of climates, whereas site-specific temporal validation can offer insights to explain population increases or decreases. It could also provide evidence for including or discounting suspicious distribution records. We suggest that both of these types of validations should be included in inter-model comparisons. In conclusion, validating pest risk models with spatial and site-specific temporal data may provide more robust and reliable results than validations with spatial data alone.

Acknowledgements

We appreciate ArcGIS technical support from Yu Takeuchi and David Christie at the Center of Integrated Pest Management in North Carolina State University. ZedX, Inc. (Bellefonte, PA), a commercial information technology (IT) company, prepared the NCEP-RTMA and NCEP-CFSR hourly weather data sets for specific years and locations in the United States and globally. Special thanks to Nicholas Manoukis, USDA-ARS and Ashley Franklin USDA-APHIS-PPQ-CPHST Raleigh for critical comments and manuscript editing, respectively. We thank Marc De Meyer, Royal Museum for Central Africa, Tervuren, Belgium for *Bactrocera invadens* data. The authors also thank Darren Kriticos, CSIRO, Australia, Richard Baker, DEFRA, UK. Gericke Cook, USDA-APHIS-PPQ-CPHST Fort Collins and two anonymous reviewers for their valuable comments on the manuscript. This project was supported by the USDA-APHIS-PPQ Pest Detection program. This material was made possible, in part, by a Cooperative Agreement from the United States Department of Agriculture's Animal and Plant Health Inspection Service (APHIS). It may not necessarily express APHIS' views.

References

- Armstrong JW, Tang JM, Wang SJ (2009) Thermal death kinetics of mediterranean, Malaysian, melon, and oriental fruit fly (Diptera: Tephritidae) eggs and third instars. *Journal of Economic Entomology* 102: 522–532. doi: 10.1603/029.102.0209
- Benjamin S, Brown JM, Manikin G, Mann G, (2007) The RTMA background – hourly downscaling of RUC data to 5-km detail. In: 23rd Conference on IIPS, January 2007, San Antonio, TX American Meteorological Society, P 1.11.
- Bradley BA (2013) Distribution models of invasive plants over-estimate potential impact. *Biological Invasions* 15: 1417–1429. doi: 10.1007/s10530-012-0380-0
- Burikam I, Sarnthoy O, Charernsom K, Kanno T, Homma H (1992) Cold temperature treatment for mangosteens infested with the oriental fruit-fly (Diptera: Tephritidae). *Journal of Economic Entomology* 85: 2298–2301. doi: 10.1093/jee/85.6.2298

- Chen P, Ye H (2007) Population dynamics of *Bactrocera dorsalis* (Diptera: Tephritidae) and analysis of factors influencing populations in Baoshanba, Yunnan, China. *Entomological Science* 10: 141–147. doi: 10.1111/j.1479-8298.2007.00208.x
- Christenson LD, Foote RH (1960) Biology of fruit flies. *Annual Review of Entomology* 5: 171–192. doi: 10.1146/annurev.en.05.010160.001131
- Clarke AR, Armstrong KF, Carmichael AE, Milne JR, Raghu S, Roderick GK, Yeates DK (2005) Invasive phytophagous pests arising through a recent tropical evolutionary radiation: The *Bactrocera dorsalis* complex of fruit flies. *Annual Review of Entomology*. 293–319. doi: 10.1146/annurev.ento.50.071803.130428
- Cornelius ML, Duan JJ, Messing RH (2000) Volatile host fruit odors as attractants for the oriental fruit fly (Diptera: Tephritidae). *Journal of Economic Entomology* 93: 93–100. doi: 10.1603/0022-0493-93.1.93
- De Lima CPF, Jessup AJ, Cruickshank L, Walsh CJ, Mansfield ER (2007) Cold disinfestation of citrus (Citrus srm.) for Mediterranean fruit fly (*Ceratitidis capitata*) and (Diptera: Tephritidae). *New Zealand Journal of Crop and Horticultural Science* 35: 39–50.
- De Pondecas MS, Manikin GS, DiMego G, Benjamin SG, Parrish DF, Purser RJ, Wu W-S, Horel JD, Myrick DT, Lin Y (2011) The real-time mesoscale analysis at NOAA’s National Centers for Environmental Prediction: Current status and development. *Weather and Forecasting* 26: 593–612. doi: 10.1175/WAF-D-10-05037.1
- de Villiers M, Hattingh V, Kriticos DJ (2013) Combining field phenological observations with distribution data to model the potential distribution of the fruit fly *Ceratitidis rosa* Karsch (Diptera: Tephritidae). *Bulletin of Entomological Research* 103: 60–73. doi: 10.1017/s0007485312000454
- Dentener PR, Alexander SM, Lester PJ, Petry RJ, Maindonald JH, McDonald RM (1996) Hot air treatment for disinfestation of lightbrown apple moth and longtailed mealy bug on persimmons. *Postharvest Biology and Technology* 8: 143–152. doi: 10.1016/0925-5214(95)00068-2
- Dile YT, Srinivasan R (2014) Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: an application in the Blue Nile River Basin. *JAWRA Journal of the American Water Resources Association* 50(5): 1226–1241. doi: 10.1111/jawr.12182
- Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40: 677–697. doi: 10.1146/annurev.ecolsys.110308.120159
- Elith J, Simpson J, Hirsch M, Burgman MA (2012) Taxonomic uncertainty and decision making for biosecurity: spatial models for myrtle/guava rust. *Australasian Plant Pathology* 42: 43–51. doi: 10.1007/s13313-012-0178-7
- Eskafi FM, Fernandez A (1990) Larval-pupal mortality of Mediterranean fruit-fly (Diptera: Tephritidae) from interaction of soil, moisture, and temperature. *Environmental Entomology* 19: 1666–1670. doi: 10.1093/ee/19.6.1666
- Fitt G, Dillon M, Hamilton J (1995) Spatial dynamics of *Helicoverpa* populations in Australia: simulation modelling and empirical studies of adult movement. *Computers and Electronics in Agriculture* 13: 177–192. doi: 10.1016/0168-1699(95)00024-X

- Flitters NE, Messenger PS (1953) Bioclimatic studies of oriental fruit fly in Hawaii. *Journal of Economic Entomology* 46: 401–403. doi: 10.1093/jee/46.3.401
- Fuka DR, Walter MT, MacAlister C, Degaetano AT, Steenhuis TS, Easton ZM (2013) Using the Climate Forecast System Reanalysis as weather input data for watershed models. *Hydrological Processes* 28(22): 5613–5623. doi:10.1002/hyp.10073
- Goergen G, Vayssieres JF, Gnanvossou D, Tindo M (2011) *Bactrocera invadens* (Diptera: Tephritidae), a New Invasive Fruit Fly Pest for the Afrotropical Region: Host Plant Range and Distribution in West and Central Africa. *Environmental Entomology* 40: 844–854. doi:10.1603/en11017
- Grout TG, Daneel JH, Mohamed SA, Ekesi S, Nderitu PW, Stephen PR, Hattingh V (2011) Cold susceptibility and disinfestation of *Bactrocera invadens* (Diptera: Tephritidae) in oranges. *Journal of Economic Entomology* 104: 1180–1188. doi: 10.1603/ec10435
- Gutierrez AP, Mills NJ, Ponti L (2010) Limits to the potential distribution of light brown apple moth in Arizona-California based on climate suitability and host plant availability. *Biological Invasions* 12: 3319–3331. doi: 10.1007/s10530-010-9725-8
- Han P, Wang X, Niu CY, Dong YC, Zhu JQ, Desneux N (2011) Population dynamics, phenology, and overwintering of *Bactrocera dorsalis* (Diptera: Tephritidae) in Hubei Province, China. *Journal of Pest Science* 84: 289–295. doi: 10.1007/s10340-011-0363-4
- Heather NW, Whitfort L, McLauchlan RL, Kopittke R (1996) Cold disinfestation of Australian mandarins against Queensland fruit fly (Diptera: Tephritidae). *Postharvest Biology and Technology* 8: 307–315. doi: 10.1016/0925-5214(96)00022-1
- Hou BH, Xie Q, Zhang RJ (2006) Depth of pupation and survival of the oriental fruit fly, *Bactrocera dorsalis* (Diptera: Tephritidae) pupae at selected soil moistures. *Applied Entomology and Zoology* 41: 515–520. doi:10.1303/aez.2006.515
- Jang EB, Nagata JT, Chan HT, Laidlaw WG (1999) Thermal death kinetics in eggs and larvae of *Bactrocera latifrons* (Diptera: Tephritidae) and comparative thermotolerance to three other tephritid fruit fly species in Hawaii. *Journal of Economic Entomology* 92: 684–690. doi: 10.1093/jee/92.3.684
- Jayanthi PDK, Verghese A (2011) Host-plant phenology and weather based forecasting models for population prediction of the oriental fruit fly, *Bactrocera dorsalis* Hendel. *Crop Protection* 30: 1557–1562. doi:10.1016/j.cropro.2011.09.002
- Jessup AJ (1992) Low-temperature storage as a quarantine treatment for table grapes infested with Queensland fruit-fly (*Bactrocera tryoni* Froggatt). *New Zealand Journal of Crop and Horticultural Science* 20: 235–239. doi: 10.1080/01140671.1992.10421921
- Jessup AJ, Delima CPF, Hood CW, Sloggett RF, Harris AM, Beckingham M (1993) Quarantine disinfestation of lemons against *Bactrocera tryoni* and *Ceratitis capitata* (Diptera: Tephritidae) using cold storage. *Journal of Economic Entomology* 86: 798–802. doi: 10.1093/jee/86.3.798
- Jessup AJ, Sloggett RF, Quinn NM (1998) Quarantine disinfestation of blueberries against *Bactrocera tryoni* (Froggatt) (Diptera: Tephritidae) by cold storage. *Journal of Economic Entomology* 91: 964–967. doi: 10.1093/jee/91.4.964
- Jose MS, Leblanc L, Geib SM, Rubinoff D (2013) An Evaluation of the Species Status of *Bactrocera invadens* and the Systematics of the *Bactrocera dorsalis* (Diptera: Tephritidae) Complex. *Annals of the Entomological Society of America* 106: 684–694. doi: 10.1603/an13017

- Kaliyan N, Carrillo MA, Morey RV, Wilcke WF, Kells SA (2007) Mortality of Indian-meal moth (Lepidoptera: Pyralidae) populations under fluctuating low temperatures: Model development and validation. *Environmental Entomology* 36: 1318–1327. doi: 10.1603/0046-225x(2007)36[1318:moimlp]2.0.co;2
- Kriticos D, Morin L, Webber B (2014) Taxonomic uncertainty in pest risks or modelling artefacts? Implications for biosecurity policy and practice. *NeoBiota* 23: 81–93. doi:10.3897/neobiota.23.7496
- Kriticos DJ, Brown JR, Maywald GF, Radford ID, Mike Nicholas D, Sutherst RW, Adkins SW (2003) SPAnDX: a process-based population dynamics model to explore management and climate change impacts on an invasive alien plant, *Acacia nilotica*. *Ecological Modelling* 163: 187–208. doi: 10.1016/S0304-3800(03)00009-7
- Legaspi BC, Legaspi JC (2010) Field-level validation of a CLIMEX model for *Cactoblastis cactorum* (Lepidoptera: Pyralidae) using estimated larval growth rates. *Environmental Entomology* 39: 368–377. doi: 10.1603/en08248
- Legates DR, McCabe GJ (1999) Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Research* 35: 233–241. doi: 10.1029/1998wr900018
- Lenzen M, Moran D, Kanemoto K, Foran B, Lobefaro L, Geschke A (2012) International trade drives biodiversity threats in developing nations. *Nature* 486: 109–112. doi: 10.1038/nature11145
- Macfadyen S, Kriticos DJ (2012) Modelling the geographical range of a species with variable life-history. *PLoS ONE* 7: e40313. doi: 10.1371/journal.pone.0040313
- Magarey R, Seem R, Russo J, Zack J, Waight K, Travis J, Oudemans P (2001) Site-specific weather information without on-site sensors. *Plant Disease* 85: 1216–1226. doi: 10.1094/PDIS.2001.85.12.1216
- Magarey RD, Borchert DM, Fowler GL, Hong SC (2014) The NCSU/APHIS Plant Pest Forecasting System (NAPPFAS^T). In: Venette RC (Ed.) *Pest Risk Modeling and Mapping for Invasive Alien Species*. CABI, Wallingford, UK.
- Magarey RD, Fowler GA, Borchert DM, Sutton TB, Colunga-Garcia M, Simpson JA (2007) NAPPFAS^T: an internet system for the weather-based mapping of plant pathogens. *Plant Disease* 91: 336–345. doi: 10.1094/PDIS-91-4-0336
- Maliva R, Missimer T (2012) *Arid Lands Water Evaluation and Management*, Environmental Science and Engineering. Springer-Verlag, Germany.
- Maywald GF, Kriticos DJ, Sutherst RW, Bottomley W (2007) DYMEX model builder version 3: user’s guide. Hearne Publishing, Melbourne. <http://blog.daum.net/dnjsgmlo3o/134>
- Monfreda C, Ramankutty N, Foley JA (2008) Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles* 22: 1–19. doi: 10.1029/2007GB002947
- Papadopoulos NT, Plant RE, Carey JR (2013) From trickle to flood: the large-scale, cryptic invasion of California by tropical fruit flies. *Proceedings of the Royal Society B: Biological Sciences* 280. doi: 10.1098/rspb.2013.1466
- Pardey P, Beddow J, Kriticos D, Hurley T, Park R, Duveiller E, Sutherst R, Burdon J, Hodson D (2013) Right-sizing stem-rust research. *Science* 340: 147–148. doi: 10.1126/science.122970

- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231–259. doi: 10.1016/j.ecolmodel.2005.03.026
- Pinkard E, Battaglia M, Bruce J, Leriche A, Kriticos D (2010a) Process-based modelling of the severity and impact of foliar pest attack on eucalypt plantation productivity under current and future climates. *Forest Ecology and Management* 259: 839–847. doi: 10.1016/j.foreco.2009.06.027
- Pinkard E, Kriticos D, Wardlaw T, Carnegie A, Leriche A (2010b) Estimating the spatiotemporal risk of disease epidemics using a bioclimatic niche model. *Ecological Modelling* 221: 2828–2838. doi: 10.1016/j.ecolmodel.2010.08.017
- Pinkard EA, Battaglia M, Bruce J, Leriche A, Kriticos DJ (2010c) Process-based modelling of the severity and impact of foliar pest attack on eucalypt plantation productivity under current and future climates. *Forest Ecology and Management* 259: 839–847. doi: 10.1016/j.foreco.2009.06.027
- Saha S, Moorthi S, Pan HL, Wu XR, Wang JD, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D, Liu HX, Stokes D, Grumbine R, Gayno G, Wang J, Hou YT, Chuang HY, Juang HMM, Sela J, Iredell M, Treadon R, Kleist D, Van Delst P, Keyser D, Derber J, Ek M, Meng J, Wei HL, Yang RQ, Lord S, Van den Dool H, Kumar A, Wang WQ, Long C, Chelliah M, Xue Y, Huang BY, Schemm JK, Ebisuzaki W, Lin R, Xie PP, Chen MY, Zhou ST, Higgins W, Zou CZ, Liu QH, Chen Y, Han Y, Cucurull L, Reynolds RW, Rutledge G, Goldberg M (2010) The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society* 91: 1015–1057. doi: 10.1175/2010bams3001.1
- Schutze MK, Aketarawong N, Amornsak W, Armstrong KF, Augustinos AA, Barr N, Bo W, Bourtzis K, Boykin LM, C  ceres C, Cameron SL, Chapman TA, Chinvinijkul S, Chomi   A, De Meyer M, Drosopoulou E, Englezou A, Ekesi S, Gariou-Papalexiou A, Geib SM, Hailstones D, Hasanuzzaman M, Haymer D, Hee AKW, Hendrichs J, Jessup A, Ji Q, Khamis FM, Krosch MN, Leblanc LUC, Mahmood K, Malacrida AR, Mavragani-Tsipidou P, Mwatawala M, Nishida R, Ono H, Reyes J, Rubinoff D, San Jose M, Shelly TE, Srikachar S, Tan KH, Thanaphum S, Haq I, Vijaysegaran S, Wee SL, Yesmin F, Zacharopoulou A, Clarke AR (2014) Synonymization of key pest species within the *Bactrocera dorsalis* species complex (Diptera: Tephritidae): taxonomic changes based on a review of 20 years of integrative morphological, molecular, cytogenetic, behavioural and chemoecological data. *Systematic Entomology* 40(2): 456–471. doi: 10.1111/syen.12113
- Sutherst RW, Maywald G, Kriticos DJ (2007) CLIMEX Version 3: User's Guide. Hearne Scientific Software Pty Ltd, Melbourne. <http://www.hearne.com.au/attachments/Climex-UserGuide3.pdf>
- Sutherst RW, Maywald GF (1985) A computerized system for matching climates in ecology. *Agriculture Ecosystems & Environment* 13: 281–299. doi: 10.1016/0167-8809(85)90016-7
- Sutherst RW, Maywald GF, Yonow T, Stevens PM (1999) CLIMEX: predicting the effects of climate on plants and animals. CSIRO Publishing, Collingwood, Australia, 88 pp.
- Travis J, Rajotte E, Bankert R, Hickey K, Hull L, Eby V, Heinemann P, Crassweller R, McClure J, Bowser T (1992) A working description of the Penn State Apple Orchard Consultant, an expert system. *Plant Disease* 76: 545–554. doi: 10.1094/PD-76-0545

- Vargas RI, Chang HB, Komura M, Kawamoto D (1987) Mortality, stadia duration, and weight-loss in 3 species of mass-reared fruit-fly pupae (Diptera, Tephritidae) held with and without vermiculite at selected relative humidities. *Journal of Economic Entomology* 80: 972–974. doi: 10.1093/jee/80.4.972
- Vargas RI, Mau RFL, Stark JD, Pinero JC, Leblanc L, Souder SK (2010) Evaluation of methyl eugenol and cue-lure traps with solid lure and insecticide dispensers for fruit fly monitoring and male annihilation in the Hawaii areawide pest management program. *Journal of Economic Entomology* 103: 409–415. doi: 10.1603/ec09299
- Vargas RI, Nishida T, Beardsley JW (1983) Distribution and abundance of *Dacus dorsalis* (Diptera: Tephritidae) in native and exotic forest areas on Kauai. *Environmental Entomology* 12: 1185–1189. doi: 10.1093/ee/12.4.1185
- Vargas RI, Stark JD, Nishida T (1990) Population dynamics, habitat preference, and seasonal distribution patterns of oriental fruit fly and melon fly (Diptera: Tephritidae) in an agricultural area. *Environmental Entomology* 19: 1820–1828.
- Vargas RI, Walsh WA, Jang EB, Armstrong JW, Kanehisa DT (1996) Survival and development of immature stages of four Hawaiian fruit flies (Diptera: Tephritidae) reared at five constant temperatures. *Annals of the Entomological Society of America* 89: 64–69. doi: 10.1093/aesa/89.1.64
- Vargas RI, Walsh WA, Kanehisa D, Stark JD, Nishida T (2000) Comparative demography of three Hawaiian fruit flies (Diptera : Tephritidae) at alternating temperatures. *Annals of the Entomological Society of America* 93: 75–81. doi: 10.1603/0013-8746(2000)093[0075:cd othf]2.0.co;2
- Waage JK, Fraser RW, Mumford JD, Cook DC, Wilby A (2005) *A New Agenda for Biosecurity*. DEFRA, London, 185 pp.
- Webber BL, Yates CJ, Le Maitre DC, Scott JK, Kriticos DJ, Ota N, McNeill A, Le Roux JJ, Midgley GF (2011) Modelling horses for novel climate courses: insights from projecting potential distributions of native and alien Australian acacias with correlative and mechanistic models. *Diversity and Distributions* 17: 978–1000. doi: 10.1111/j.1472-4642.2011.00811.x
- Willmott CJ (1981) On the validation of models. *Physical Geography* 2: 184–194
- Xie Q, Hou BH, Zhang RJ (2008) Thermal responses of oriental fruit fly (Diptera: tephritidae) late third instars: Mortality, puparial morphology, and adult emergence. *Journal of Economic Entomology* 101: 736–741. doi: 10.1603/0022-0493(2008)101[736:trooff]2.0.co;2
- Xie Q, Zhang RJ (2007) Responses of oriental fruit fly (Diptera: Tephritidae) third instars to desiccation and immersion. *Journal of Agricultural and Urban Entomology* 24: 1–11. doi: 10.3954/1523-5475-24.1.1
- Yang PJ, Carey JR, Dowell RV (1994) Temperature influences on the development and demography of *Bactrocera dorsalis* (Diptera: Tephritidae) in China. *Environmental Entomology* 23: 971–974. doi: 10.1093/ee/23.4.971
- Yonow T, Zalucki MP, Sutherst RW, Dominiak BC, Maywald GF, Maelzer DA, Kriticos DJ (2004) Modelling the population dynamics of the Queensland fruit fly, *Bactrocera (Dacus) tryoni*: a cohort-based approach incorporating the effects of weather. *Ecological Modelling* 173: 9–30. doi: 10.1016/s0304-3800(03)00306-5

Supplementary material 1

Table S1

Authors: Seung Cheon Hong, Roger D. Magarey, Daniel M. Borchert, Roger I. Vargas, Steven K. Souder

Data type: distribution data

Explanation note: List of references for oriental fruit fly (*B. dorsalis*) distribution.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 2

Table S2

Authors: Seung Cheon Hong, Roger D. Magarey, Daniel M. Borchert, Roger I. Vargas, Steven K. Souder

Data type: distribution data

Explanation note: List of references for *B. invadens* distribution (Courtesy to Marc De Meyer, Royal Museum for Central Africa, Tervuren, Belgium).

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 3

Table S3

Authors: Seung Cheon Hong, Roger D. Magarey, Daniel M. Borchert, Roger I. Vargas, Steven K. Souder

Data type: specimens data

Explanation note: Summary statistics of low and high temperature mortality of *Bactrocera dorsalis*.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 4

Figure S1

Authors: Seung Cheon Hong, Roger D. Magarey, Daniel M. Borchert, Roger I. Vargas, Steven K. Souder

Data type: specimens data

Explanation note: Fit of low temperature mortality function (line) to observations (diamonds).

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 5

Figure S2

Authors: Seung Cheon Hong, Roger D. Magarey, Daniel M. Borchert, Roger I. Vargas, Steven K. Souder

Data type: specimens data

Explanation note: Fit of high temperature mortality function (line) to observations (diamonds).

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.